STATISTICS Ph. D. COMPREHENSIVE EXAM

January, 2002

General Instructions: Write your ID number on your answer sheets. Do not put your name on any of your answer sheets. All solutions should be rigorously explained. Show all work. If you cannot solve a problem, at least explain what is this problem about and what approach may lead to its solution.

Reminder: All assertions should be rigorously proved.

1. For a linear model $Y = X\beta + e$, $E(e) = 0$, $Cov(e) = \sigma^2 I$, the residuals are $\hat{e} = (I - M)Y$ where M is the perpedicular projection operator (PPO) onto the $C(X)$. Find the following:

   (a) $E(\hat{e})$.

   (b) $Cov(\hat{e})$.

   (c) $Cov(\hat{e}, MY)$.

   (d) $E(\hat{e}'\hat{e})$.

2. Consider the one-way analysis of variance model

   $$y_{ij} = \mu + \alpha_i + e_{ij}, \quad i = 1, 2, 3; \quad j = 1, ..., N_i; \quad (N_1, N_2, N_3) = (3, 2, 1),$$

   where the e_{ij}'s are independent $N(0, \sigma^2)$. You may need some of the following additional information:

   $$X'X = \begin{bmatrix} 6 & 3 & 2 & 1 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad (X'X)^{-} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/3 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

   $$(X'X)^{-}X' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\text{and } X(X'X)^{-}X' = \begin{bmatrix} 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 1/3 & 1/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

(a) Write out the design matrix X.

(b) Show that $\alpha_1 - \alpha_3$ is estimable.

(c) Find $\hat{\beta}$.

(d) Find the least squares estimate of $\alpha_1 - \alpha_3$.

(e) Show that for this model, $\lambda'\beta = (0, c_1, c_2, c_3)\beta$ is estimable iff it is a contrast. (The c_i are constants).

3. Let $X_1, X_2, \ldots, X_n$ be *iid* with distribution $F(x - \theta)$. Suppose that $F(0) = 1/2$ and that at zero F has a density $f(0) > 0$.

(a) Find the asymptotic distribution of the sample median.

(b) Compare asymptotic properties of the sample mean and sample median estimates for the standard normal distribution; i.e. $f(x) = \dfrac{1}{\sqrt{2\pi}} \exp(-x^2/2)$.

4. Consider the Bayesian approach to solving a problem of testing m simple hypotheses.

(a) Prove that Bayes error of any test cannot be smaller than $1 - E\{\max_{i \in \{1,2,\ldots,m\}} \pi(i|X)\}$ where $\pi(i|X)$ is the posterior probability of hypothesis number i given observation X.

(b) Suggest a Bayes test (and prove that this test is a Bayes test). Is it always unique?

(c) Give an example where a Bayes test always rejects one of the hypotheses.

5. Let $X_1, \ldots, X_n \overset{iid}{\sim}$ Bernoulli(θ).

(a) Find the distribution of $W_k = \Pi_{i=1}^{k} X_i$ for $k = 1, \ldots, n$. What is $E(W_k)$?

(b) Find the UMVUE of θ^k for $k = 1, \ldots, n$.

(c) Find the UMVUE of the polynomial $a_0 + a_1\theta + a_2\theta^2 + \cdots a_n\theta^n$, where the a_j are fixed, known constants.

6. Let X_{ij} be distributed independently from a $N(\mu_i, \sigma_i^2)$ distribution, where $j = 1, 2, \ldots, n_i$ and $i = 1, 2, \ldots, k$ (i.e. k groups with sample sizes $n_1, n_2, \ldots, n_k$). Define $s_i^2 = \sum_{j=1}^{n_i}(X_{ij} - \bar{X}_i)^2/(n_i - 1)$, where $\bar{X}_i$ is the mean for the i^{th} group.

   (a) Show that $\sqrt{n_i}\{\log_e(s_i^2) - \log_e(\sigma_i^2)\} \xrightarrow{d} N(0, \nu)$ as $n_i \to \infty$, and give an expression for ν.

   (b) Define $Y' = (Y_1, Y_2, \ldots, Y_k)'$ where $Y_i = \log_e(s_i^2)$ and set $\theta' = (\theta_1, \theta_2, \ldots, \theta_k)'$ where $\theta_i = \log_e(\sigma_i^2)$. In large samples, what is the approximate distribution of Y?

   (c) Suppose $z_i' = (z_{i1}, z_{i2}, \ldots, z_{ip})'$ $(p \leq k)$ is a set of fixed covariate values associated with the i^{th} group and that

$$\theta_i = \log_e(\sigma_i^2) = z_i'\beta$$

   for some parameter vector β. Using part (b), suggest, and derive, an estimator for β, say $\hat{\beta}$. Present some distributional properties of $\hat{\beta}$.

   (d) Derive a test statistic for testing $H_0 : H\beta = 0$ where H is a fixed $r \times p$ matrix $(r < p)$. Be clear about how the test is performed.